in the α-helical conformation absorbs between 1650 and 1660 cm^{-1}, but for the β-extended conformation of polypeptides this band is located near 1630 cm^{-1}. Miyazawa *et al.*[2] carried out normal coordinate calculations on model compounds which showed that three strong infrared bands of proteins, called amide I, II, and III, involve coupled C-O stretching, C-N stretching, and N-H bending modes of the CONH group. They also carried out calculations showing that coupling between neighboring peptide groups occurs both through the α-carbon atoms and across hydrogen bonds.[3] The theory of Miyazawa *et al.* was modified in 1962 by Krimm[4] and again in 1972 by Krimm and Abe.[5] Reviews have been written by Kauzmann[6] and by Schellman and Schellman.[7] More recently Fourier transform infrared (FT-IR) studies of proteins in aqueous solution were reported by Koenig and Tabb.[8]

The conformation of most naturally occurring proteins can be described as being comprised of differing proportions of α-helix, extended β-strands which interact with other β-strands through hydrogen bonds, and turns, in which the polypeptide chain reverses direction, bending back on itself.[9] Infrared absorption bands due to each type of conformation are too broad and overlap too extensively to permit anything but unresolved features to be observed, making it very difficult to glean detailed conformational information. Timasheff and coworkers[10] have had some limited success with protein spectra observed in deuterium oxide solution, but their inferences were frequently made from ill-resolved shoulders of broad bands and the results were therefore tentative.

One method of reducing the width of the component bands is through the use of Fourier self-deconvolution (FSD).[11–13] In this procedure, the Fourier transform of the spectrum (stored linear in absorbance) is computed. The Fourier domain array is multiplied by an exponential weighting function, $\exp(\gamma' x)$, and the reverse transform is computed, often after truncation of the array. The greater the exponent of the weighting function, the narrower are the resulting bands in the spectrum. Two factors prevent the width of the bands from being reduced beyond a certain limit. First, the signal-to-noise ratio (SNR) of the spectrum is degraded severely on FSD and, second, even in noise-free spectra artifacts can appear in the form of side-lobes if the original spectrum was not measured at a sufficiently high resolution, if the Fourier domain array was truncated too early, or if the value selected for γ' (assumed half-width of the component bands) was too large. Griffiths and Yang[14] have developed procedures to aid in the selection of the exponent and the degree of truncation, giving the optimum combination of SNR, resolution enhancement, and side-lobe supression. These techniques have now been applied to the infrared spectra of solid proteins and protein solutions in D_2O. The results indicate that FSD should enable quantitative estimates of the contributions of each conformation to be calculated from the infrared spectra of proteins.

EXPERIMENTAL

Diffuse reflectance (DR) infrared spectra of solid β-lactoglobulin and cytochrome *c* were measured with the use of a Digilab interferometer at the University of California (UCR) with the optical configuration previously reported by Fuller and Griffiths.[15,16] A Globar source, and a medium-range ($\nu_{min} = 600$ cm^{-1}) mercury-cadmium-telluride (MCT) detector were employed for these DR measurements, and spectra were computed on the data system of a Digilab FTS-20 spectrometer. Each protein was ground with KCl for one minute at a concentration of 10% (W/W). Spectra were measured at a nominal resolution of 4 cm^{-1}, and 2000 scans were signal-averaged (about 12 min measurement time) in order that an SNR sufficiently high for the reproducible application of FSD techniques could be obtained. No apodization was applied prior to Fourier transformation. The DR spectrum of myoglobin (5% W/W in KCl) was measured in a like manner by K. Krishnan of the Digilab Division, Bio-Rad Laboratories, Cambridge, MA, but with the Digilab DR accessory and a long-range ($\nu_{min} =$ 450 cm^{-1}) MCT detector. The latter spectrum consisted of 1000 signal-averaged scans.

Transmission FT-IR spectra of proteins in the solid state as both KBr pellets and as mulls were obtained at the U.S. Department of Agriculture (USDA) at 2 cm^{-1} resolution on a Nicolet 7199 spectrometer (Nicolet Instrument Corporation, Madison, WI) equipped with a Globar source, a Ge/KBr beamsplitter, and a wide-range MCT detector ($\nu_{min} = 350$ cm^{-1}). We prepared KBr discs (7 mm diameter) by grinding about 0.4 mg of each protein with approximately 80 mg of KBr and pressing at ~12,000 psi for 2–3 min with a Harshaw Quick Press (Harshaw Chemical Co., Solon, OH). We prepared halocarbon mulls by grinding approximately 20 mg of each protein with one drop of Fluorolube. Each spectrum was measured by the signal-averaging of 4000 scans; interferograms were apodized with the Happ-Genzel function prior to Fourier transformation.

Spectra of proteins dissolved in D_2O were obtained at both USDA and UCR. At UCR proteins were dissolved in D_2O at a concentration of 2.5% (W/V) and held in a 50-μm BaF_2 cell. We measured spectra on a Nicolet 60-SX spectrometer, at a resolution 4 cm^{-1}, by averaging 2500 scans taken at a very high velocity (80 kHz data acquisition); the total measurement time was 7 min. Deconvolution was performed by the employment of the same parameters used for the DR spectra, so that a direct comparison of the results could be made.

At USDA the protein concentration used was about 5% (W/V). The CaF_2 cell had a 75-μm pathlength. All spectra were obtained at pD 7, where most side-chain carboxyl groups are ionized. We prepared solutions 48 h before measuring the spectra to assure total H, D exchange. The disappearance of the amide II band in the 1550 cm^{-1} region (primarily due to N-H bending) verified that isotopic exchange was complete. The background absorption caused by D_2O was subtracted in the usual manner. The instrumental parameters were the same as for the solid state spectra.

The average half-width, γ', used in the computations performed at UCR and Digilab was 7.5 cm^{-1}, while at USDA a γ' value of 6.5 cm^{-1} was employed. The percentage of the Fourier domain used in the computation was 48% at UCR and Digilab, and 30% at USDA. Selection of the optimum value of γ' is difficult if bands

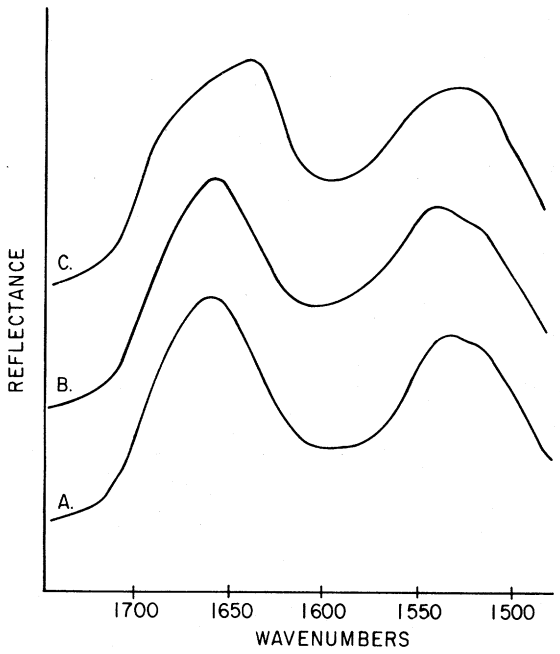

FIG. 1. Original diffuse reflectance spectra of (A) myoglobin, (B) cytochrome *c*, and (C) β-lactoglobulin.

with different half-width must be resolved. The best value for most spectral features may result in overdeconvolution of some very sharp bands (for instance, the 1515 cm^{-1} band of the proteins in Figs. 4–5). The parameters used at UCR were found with the use of the method of Yang and Griffiths,[14] to ensure that side-lobes are completely absent. The parameters used at USDA were even more conservative. All proteins except the β-lactoglobulin A used at USDA were obtained from the Sigma Chemical Co., St. Louis, MO. The latter protein was isolated and purified by Dr. H. F. Farrell, Jr., of the USDA laboratory, as previously described.[17] The D_2O was at least 99.8 atom% D.

RESULTS AND DISCUSSION

Diffuse Reflectance Spectra. The original DR spectra are shown in Fig. 1. Proteins known to have a high α-helix content,[18] myoglobin and cytochrome *c*, are shown in Fig. 1A and 1B, while a protein with a lower α-helix content and a high percentage of β-strands,[10] β-lactoglobulin, is shown in Fig. 1C. The spectra of the same proteins after deconvolution are given in Fig. 2. While the original spectra show only broad bands centered around 1660 and 1530 cm^{-1}, the deconvolved spectra exhibit a number of component bands which can be associated with different conformations.

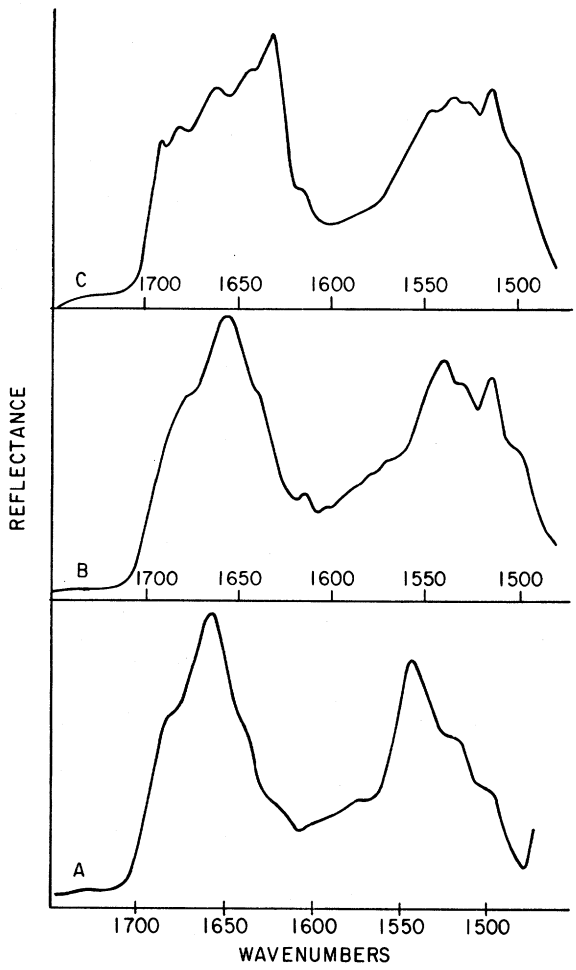

FIG. 2. Diffuse reflectance spectra after deconvolution of (A) myoglobin, (B) cytochrome *c*, and (C) β-lactoglobulin.

Over the past several years extensive normal coordinate calculations have been carried out by Krimm and his associates for the assignment of infrared absorption bands to the α-helix,[19] to the β-structure,[20] and to β-turns.[21,22] We base our assignments on these calculations

TABLE I. Amide I and amide II frequencies (cm^{-1}) observed after deconvolution of diffuse reflectance spectra.

	Myoglobin	Cytochrome *c*	β-Lactoglobulin
Amide I			
α-helix	1656	1654	1653
β-strands	...	...	1632, 1674
Turns[a]	1638, 1681	1639, 1680	1645, 1662, 1682
Amide II			
α-helix	1543	1539	...
β-strands	...	...	1548
Unassigned	...	1533	1526, 1535

[a] Tentative assignments.

TABLE II. Comparison of bands (cm^{-1}) observed in the amide I region of the deconvolved spectrum of β-lactoglobulin with the use of different sampling techniques.[a]

Technique:	Diffuse reflectance	KBr disk	Halocarbon mull	Deuterium oxide solution	
Measured at:	UCR	USDA	USDA	UCR	USDA
	1693	1695	1694	1693	1693
	1682	1684	1684	1682	(1683)
	1674	(1673)	(1672)	1679	1679
	1662	1662	1660	1664	(1664)
	1653	1652	...	...	...
	(1645)	(1644)	1645	1648	1648
	1632	1630	1628	1634	1634
	...	...	...	1624	1624
	1617	1617	1616	1615	1615

[a] Shoulders are given in parentheses.

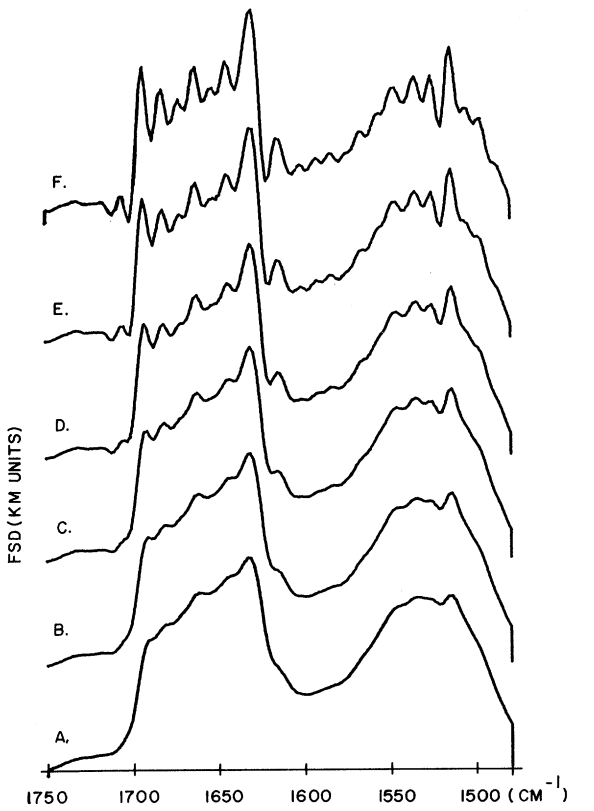

FIG. 3. Diffuse reflectance spectra after deconvolution of β-lactoglobulin for increasing values of γ': (A) 5.5 cm^{-1}, (B) 6.5 cm^{-1}, (C) 7.5 cm^{-1}, (D) 8.5 cm^{-1}, (E) 9.5 cm^{-1}, (F) 10.5 cm^{-1}.

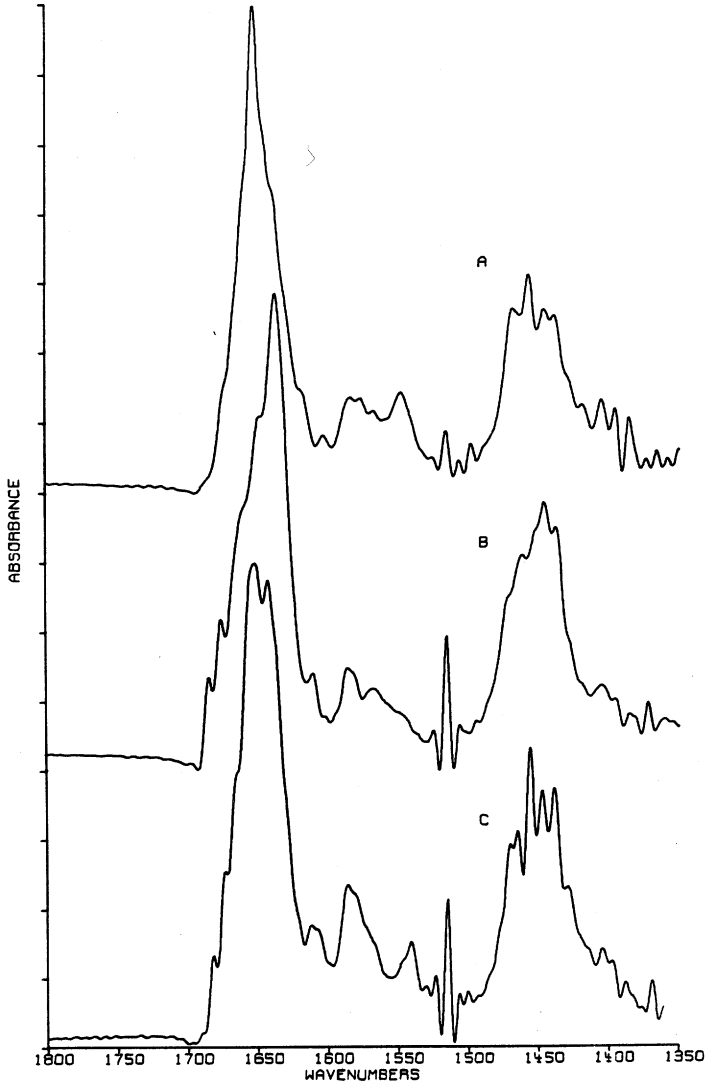

FIG. 4. Deconvolved spectra of D_2O solutions of proteins with known structure: (A) hemoglobin, (B) ribonuclease A, and (C) lysozyme.

and on empirical correlations as they have accumulated over the years.[1,3,4,8,10,19–22] In the amide I region, the α-helix absorbs around 1650–1660 cm^{-1}, while the β-strands absorb at 1630–1640 cm^{-1} (strong) and 1670–1690 cm^{-1} (weak). There is very little empirical information for turns. Calculations indicate that a series of weak bands should occur in the 1660 to 1690 cm^{-1} range,[21] in general agreement with second-derivative Fourier transform infrared spectra of proteins in D_2O solution.[23] Nonetheless, some weak bands in the 1630 to 1660 cm^{-1} range could also be caused by turns.[22] An "unordered" structure in the solid state was previously associated with bands around 1660 cm^{-1},[3,4] but this correlation appears to be unsound.[5] It should be emphasized that the α-helical and β-strand segments in globular proteins are frequently very short (three to ten residues) and distorted.[18] Correlations and calculations based on more regular models, such as synthetic polypeptides,[3–5] therefore apply only to a limited degree. A tentative assignment of the amide I and amide II components, as observed in DR spectra after deconvolution, is given in Table I.

In some cases the assignment to β-strands or turns is uncertain, because the characteristic frequencies seem to overlap.[20–23] The very sharp and stable band close to 1515 cm^{-1}, which does not shift with deuteration, is probably associated with tyrosine residues[24] and is omitted in Table I. The frequency differences between the various components of the amide II band are evidently small and not well suited for structural studies. The same conclusion emerges from Koenig's work in H_2O solution.[8]

Some of the observed bands are quite weak and could conceivably be caused by noise, side-lobes, or water vapor. Indeed, the effect of increasing the value of γ' is to increase artifacts caused by all three factors. Figure 3 illustrates the increased resolution and side-lobe amplitude in the DR spectrum of β-lactoglobulin with increasing γ'. To eliminate errors caused by noise, we repeated each measurement three times. To avoid artifacts caused by overdeconvolution, we kept our γ' value low. In Fig. 3C ($\gamma' = 7.5$) no bands are observed which are not already present at $\gamma' = 5.5$, where no distortion occurs as judged by the 1560 to 1630 cm^{-1} region. In view of the increased resolution obtained with $\gamma' = 7.5$ cm^{-1}, this value was employed for the deconvolution of all DR spectra. Absence of water vapor lines was confirmed by a detailed study of the pertinent frequencies. We therefore conclude that our observed features are real and truly represent deconvolved absorption spectra of the investigated proteins.

Disks and Mulls. Spectra of β-lactoglobulin obtained at USDA after the samples were prepared as KBr disks or halocarbon mulls were remarkably similar to the dif-

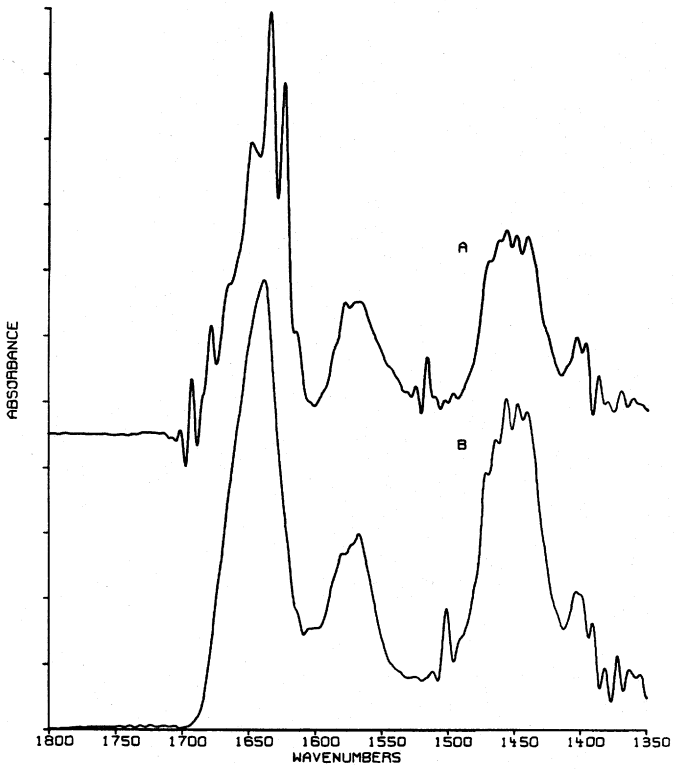

FIG. 5. Deconvolved spectra of β-lactoglobulin A in D_2O solution as a function of pD: (A) native protein at pD 7, (B) denatured protein at pD 13.

fuse reflectance spectra measured at UCR even after deconvolution. As an example, the wavenumbers of bands observed after self-deconvolution in the amide I region of the spectrum of β-lactoglobulin sampled by means of four different techniques are summarized in Table II.

Deuterium Oxide Solutions. There are three main reasons why protein spectra should be measured in aqueous environment: (1) this is the natural state of globular proteins; (2) it permits a study of structural changes, if any, accompanying protein dissolution; and (3) solution spectra are easier to analyze in a quantitative manner because of known sample thickness and the applicability of Beer's law. Because water absorbs very strongly in the region of the amide I frequency, studies are frequently carried out in D_2O solution.[10,21] When proteins are dissolved in D_2O, the amide I and II bands are usually renamed as the amide I′ and II′ bands, respectively.[2,3] Band assignments for the amide II′ band in D_2O solution may not be used for structure correlations because of overlap with the HOD bending frequency (HOD is always present in small amounts) and with side-chain modes.[21] One notable difference between solid state and solution spectra involves the "unordered" substructure. In deuterium oxide these amino acid residues are hydrogen-bonded to the solvent (rather than to other peptide groups) and produce a band around 1640–1643 cm^{-1}.[10,21]

Spectra after deconvolution of three proteins of known structure—hemoglobin, ribonuclease A, and lysozyme—are shown in Fig. 4; their amide I′ frequencies are summarized in Table III. The original spectra are not shown. They are similar to the ones in Fig. 1, but the broad, unresolved amide I′ and II′ bands now center around 1650 and 1450 cm^{-1}, respectively. The spectrum of hemoglobin after deconvolution exhibits a strong, sharp band at 1652 cm^{-1}, clearly associated with the α-helix. This protein is about 80% α-helical and has no β-structure.[18] The weak shoulders close to 1675, 1662, and 1640 cm^{-1} can thus be associated with turns.

Ribonuclease A has an α-helix content of about 22% and a β-strand content of about 46%, based on the classification of Levitt and Greer.[18] The strong 1637 cm^{-1} band can be associated with the β-segments and the much weaker 1649 cm^{-1} band with the α-helix. The weak band at 1676 cm^{-1} is assigned to the second β-structure band by comparison with ribonuclease S,[25] which has a very high β-structure content and few turns.[18] The remaining weak bands at 1662 and 1685 cm^{-1} are associated with turns.

Lysozyme has strong bands at 1652 (α-helix) and 1642 cm^{-1} (hydrated peptide groups) and weak bands at 1683, 1674, and 1666 cm^{-1} (attributable to turns). This protein has about 45% α-helix and 19% β-structure.[18] The β-segments, however, are very short and distorted (3+3+6+4+3 residues)[18] and evidently do not give rise to a typical β-structure band around 1630–1640 cm^{-1}. These short segments can form no regular sheets and some of them might be hydrated, contributing to the absorption at 1642 cm^{-1}.

Figure 5 shows spectra of β-lactoglobulin after deconvolution. The detailed structure of this protein is not known, but circular dichroism studies suggest an α-helix content of about 20% and a β-structure content of about 40–50%.[26] The features at 1634 and 1679 cm^{-1} are assigned to β-strands, the band at 1649 cm^{-1} to the α-helix, and the bands at 1664 and 1693 cm^{-1} to turns. An unusual, sharp band at 1624 cm^{-1} (not observed in the solid) could be associated with a special kind of β-strand or turn. This band is also present in concanavalin A,[25] but evidently not in ribonuclease A, both of which have a very high β-content.[18] Denatured β-lactoglobulin shown

TABLE III. Amide I′ frequencies (cm^{-1}) in deuterium oxide solution.[a]

Protein	α-helix	β-strands		Turns				Hydrated
Hemoglobin	1652	···	···	[1640]	1662	1675	···	···
Ribonuclease A	1649	1636	1676	···	1662	···	1685	···
Lysozyme	1652	···	···	···	1666	1674	1683	1642
β-lactoglobulin (native, pD 7)	1649	1634	1679	[1624]	1664	···	1693	···
β-lactoglobulin (denatured, pD 13)	···	···	···	···	···	···	···	1640

[a] Uncertain assignments are given in square brackets.

in Fig. 5B has a single strong band about 1640 cm^{-1}, associated with solvent bonded peptide groups of the now "unordered" protein. Note that neither native β-lactoglobulin nor ribonuclease A has such a band, indicating that the segments of these proteins called "unordered" by an earlier nomenclature[3,4,10] are most probably "turns."

The bands below about 1600 cm^{-1} observed in all protein spectra after deconvolution can be associated with aromatic, carboxylate, methylene, and methyl groups in the side-chains.[21,24] While more work is needed for detailed assignments, it is evident that deconvolution should also be useful in studies of the side-chains, as well as for further investigation into the conformation of proteins.

Reproducibility of Deconvoluted Spectra. It has sometimes been suggested that the extent of deconvolution is somewhat arbitrary and that different workers will usually arrive at different results. In the last column of Table II, the wavenumbers of bands in the deconvolved spectra of β-lactoglobulin in D_2O solution independently measured at UCR and at USDA are compared. The frequency agreement is very good throughout and the intensity patterns were the same. It is comforting to observe that spectra measured and manipulated independently at two different laboratories on samples of different origin give such similar results. Our confidence is further increased by good agreement with second derivative spectra.[21] It was found during the investigation of DR spectra that if protein concentrations larger than 10% were examined, bands could become distorted near the band maximum and lead to the appearance of artifacts upon deconvolution. Care should be taken to ensure that bands are not saturated and that the shape of all bands in the original spectra are as close to Lorentzian as possible.

CONCLUSIONS

The feasibility of using diffuse reflectance (DR) spectrometry as a technique for measuring the infrared spectra of solid proteins has been demonstrated. Structural information derived from DR spectra and from solution spectra after deconvolution is internally consistent and in agreement with studies by other methods. Fourier self-deconvolution (FSD) provides unique, new information on "turns" and on hydrated amino acid residues. No such data have, to our knowledge, yet been obtained by other spectroscopic techniques such as circular dichroism or Raman spectroscopy. Deconvolution of the infrared spectra of proteins should also be useful for investigating side-chain structure and for the study of conformational changes of proteins in aqueous solution. Accurate quantitative data are still hard to obtain. We are investigating the feasibility of using both factor analysis and curve-fitting techniques in conjunction with FSD. The results will be reported in subsequent papers. Finally, independent measurements at two laboratories (UCR and USDA) have yielded very similar results, indicating that Fourier self-deconvolution of high-quality infrared spectra is less susceptible to error or bias than has been suggested in the past.

ACKNOWLEDGMENT

The authors are grateful to Dr. K. Krishnan of the Digilab Division of Bio-Rad Laboratories, Cambridge, MA, for obtaining the DR spectrum of myoglobin and for his helpful discussions concerning this project. D.M.B. and H.S. thank Janine Brouillette of USDA for technical assistance with sample preparation and for obtaining FT-IR spectra of the proteins in D_2O solution.

1. A. Elliott and E. J. Ambrose, Nature **4206,** 921 (1950).
2. T. Miyazawa, T. Shimanouchi, and S. Mizushima, J. Chem. Phys. **24,** 408 (1956).
3. T. Miyazawa, J. Chem. Phys. **32,** 1647 (1960).
4. S. Krimm, J. Mol. Biol. **4,** 528 (1962).
5. S. Krimm and Y. Abe, Proc. Nat. Acad. Sci. USA **69,** 2788 (1972).
6. W. Kauzmann, Ann. Rev. Phys. Chem. **8,** 413 (1975).
7. J. A. Schellman and C. Schellman, in *The Proteins,* H. Neurath, Ed. (Academic Press, New York, 1964), Vol. 2, 2nd ed., pp. 1–137.
8. J. L. Koenig and D. L. Tabb, in *Analytical Application of FT-IR to Molecular and Biological Systems,* J. R. Durig, Ed. (D. Reidel Publishing Co., Dordrecht, Netherlands, 1980), pp. 241–255.
9. J. S. Richardson, Adv. Prot. Chem. **34,** 167 (1981).
10. S. N. Timasheff, H. Susi, and L. Stevens, J. Biol. Chem. **242,** 5467 (1967).
11. J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, and D. G. Cameron, Anal. Chem. **53,** 145 (1981).
12. J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, and D. G. Cameron, Appl. Spectrosc. **35,** 271 (1981).
13. J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, and D. G. Cameron, Appl. Opt. **20,** 1866 (1981).
14. P. R. Griffiths and W.-J. Yang, Computer Enhanced Spectroscopy **1,** 157 (1984).
15. M. P. Fuller and P. R. Griffiths, Amer. Lab. **10(10),** 69 (1978).
16. M. P. Fuller and P. R. Griffiths, Anal. Chem. **50,** 1906 (1978).
17. R. Aschaffenburg and J. Drewry, Nature **176,** 218 (1955).
18. M. Levitt and J. Greer, J. Mol. Biol. **114,** 181 (1977).
19. J. F. Rabolt and S. Krimm, Macromolecules **10,** 1065 (1977).
20. A. M. Dwivedi and S. Krimm, Macromolecules **15,** 186 (1982).
21. S. Krimm and J. Bandekar, Biopolymers **19,** 1 (1980).
22. J. Bandekar and S. Krimm, Biopolymers **19,** 31 (1980).
23. H. Susi and D. M. Byler, Biochem. Biophys. Res. Comm. **115,** 391 (1983).
24. Y. N. Chirgadze, O. V. Fedorov, and N. P. Trushina, Biopolymers **14,** 679 1975).
25. D. M. Byler and H. Susi, unpublished results (1983).
26. S. N. Timasheff, R. Townend and L. Mescanti, J. Biol. Chem. **241,** 1863 (1966).

Protein Conformation by Infrared Spectroscopy: Resolution Enhancement by Fourier Self-Deconvolution

WANG-JIH YANG, PETER R. GRIFFITHS, D. MICHAEL BYLER,* and HEINO SUSI

Department of Chemistry, Ohio University, Athens, Ohio 45701 (W.-J.Y.); Department of Chemistry, University of California, Riverside, California 92521 (P.R.G.); and U.S. Department of Agriculture, Eastern Regional Research Center, 600 East Mermaid Lane, Philadelphia, Pennsylvania 19118 (D.M.B., H.S.)

Fourier self-deconvolution (FSD) has been employed to enhance the resolution of the infrared spectra of proteins in the solid state and in D_2O solution. The feasibility of using diffuse reflectance spectrometry for measuring the infrared spectra of solid proteins has been demonstrated. FSD permits inherently broad absorption bands to be resolved into distinct peaks which can be associated with specific protein secondary structures. Because the areas of the resolved peaks are the same as the areas of the previously unidentifiable components, this new method should enable quantitative estimates of the proportion of each conformation in a protein to be calculated.

Index Headings: Protein confirmation; Infrared spectroscopy; FT-IR; Deconvolution; Diffuse reflectance; Resolution enhancement.

INTRODUCTION

It has been recognized for over thirty years that the infrared absorption bands associated with amide groups of proteins and polypeptides yield information on the conformation of these biopolymers. In 1950, Elliott and Ambrose[1] showed that the amide I band of polypeptides

Received 18 June 1984.
* Author to whom correspondence should be addressed.

0003-7028/85/3902-0282$2.00/0